

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-101637

(43)Date of publication of application : 07.04.2000

(51)Int.Cl.

H04L 12/56

(21)Application number : 10-267673

(71)Applicant : FUJITSU LTD

(22)Date of filing : 22.09.1998

(72)Inventor : KYO EN

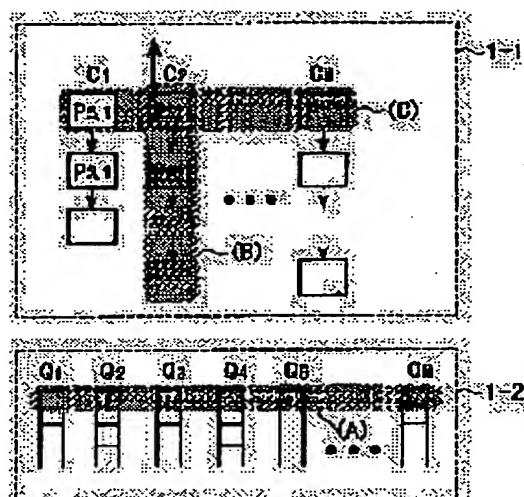
NAKAATO AKIRA

(54) PACKET TRANSFER CONTROLLER AND SCHEDULING METHOD THEREFOR

(57)Abstract:

PROBLEM TO BE SOLVED: To enable high-speed packet transfer remarkably reducing the quantity of calculation for selecting the packet to be most preferentially transferred concerning a packet transfer controller in a multi-queue configuration provided with plural queues for storing packets and a scheduling method therefor.

SOLUTION: When transferring the leading packet of N pieces of queues (Q1-QN) 1-2 corresponding to respective flows while selecting the leading packet of earliest transfer end schedule time, a scheduler 1-1 clusters the leading packets of low dependence on the relative rank fluctuation in the transfer end schedule time in respect to the fluctuation of entire reserved bands of back log queue, the packet of the first rank with the earliest transfer end schedule time among clusters C1-CN is held and at the time of transfer output, the transfer end schedule time is calculated only to the packet of the first rank among respective clusters utilizing again the output priority of respective clusters calculated in the past as it is so that the most earliest packet can be selected as the most preferential packet.



LEGAL STATUS

[Date of request for examination] 20.11.2000

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3306705

[Date of registration] 17.05.2002

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's]

decision of rejection]

[Date of extinction of right]

Copyright (C) 1998,2003 Japan Patent Office

(19)日本国特許庁(JP)

(12) 公開特許公報(A)

(11)特許出願公開番号

特開2000-101637

(P2000-101637A)

(43)公開日 平成12年4月7日(2000.4.7)

(51)Int.Cl.

H04L 12/56

識別記号

F1

H04L 11/20

テーマコード(参考)

102A 5K030

審査請求 未請求 請求項の数16 OL (全14頁)

(21)出願番号

特願平10-267673

(22)出願日

平成10年9月22日(1998.9.22)

(71)出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番
1号

(72)発明者 許 炎

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(72)発明者 中後 明

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(74)代理人 100072833

弁理士 柏谷 昭司 (外2名)

Fターム(参考) 5K030 GA01 GA08 HA08 KA02 LC09

MB11 MB13

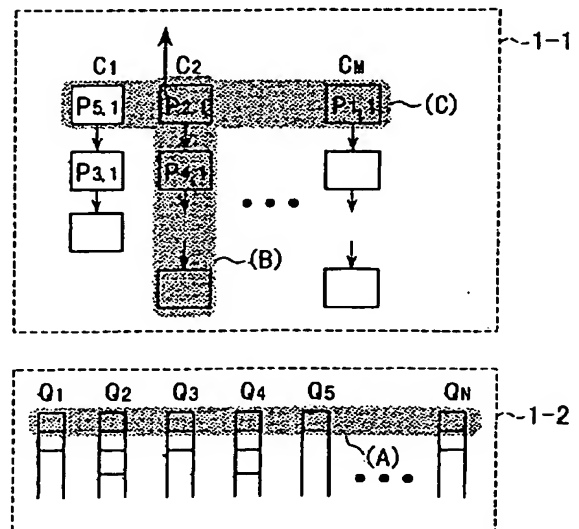
(54)【発明の名称】 パケット転送制御装置及びそのスケジューリング方法

(57)【要約】

【課題】 パケットを格納する複数のキューを備えたマルチキュー構成のパケット転送制御装置及びそのスケジューリング方法に関し、最優先に転送するパケットを選択するための計算量を大幅に削減し、高速パケット転送処理を可能にする。

【解決手段】 スケジューラ1-1は、各フロー対応のN個のキュー(Q₁~Q_N)1-2の先頭パケットに対して、転送終了予定時刻の最も早い先頭パケットを選択して転送する処理を行うとき、バックログキューの予約バンド幅全体の変動に対する転送終了予定時刻の相対的な順位変動の依存性の低い先頭パケット同士をクラスタ化し、該クラスタC₁~C_M内で転送終了予定時刻の最も早い第1順位パケットを保持し、転送出力時に、過去に計算した各クラスタ内の出力優先順位をそのまま再利用し、クラスタ間において、各クラスタの第1順位パケットのみに対して転送終了予定時刻の計算を行なって、その最も早いパケットを最優先パケットとして選出する。

本発明のパケット転送制御の原理説明図



【特許請求の範囲】

【請求項1】 予め個々に定められた予約バンド幅が割り当てられる複数のキューに格納されたパケットに対して、各キューの先頭パケットに、転送終了予定時刻の早い順に転送順序の重み付けを行うとともに、転送すべきパケットが格納されたアクティブキューの全てに、リソース（バンド幅）を割り当てて転送制御を行うスケジューラを少なくとも1つ備えたパケット転送制御装置において、前記スケジューラは、

各キューの先頭パケットのパケット長と前記予約バンド幅との比（以下、リソース占有予定時間という。）に応じて、各キューの先頭パケットをクラスタに分けるクラスタ分け手段と、

前記クラスタ内のパケットのうちで転送終了予定時刻の最も早いパケットを第1順位パケットとして、各クラスタ毎に保持するクラスタ管理手段と、

前記クラスタ内にパケットの削除又は追加があったときに、該クラスタに対して転送終了予定時刻の最も早いパケットを第1順位パケットとして選出し、前記クラスタ管理手段に保持させるクラスタ内選択手段と、

前記各クラスタの第1順位パケットの中から、転送終了予定時刻の最も早いパケットを最優先転送パケットとして選出するクラスタ間選択手段とを備えたことを特徴とするパケット転送制御装置。

【請求項2】 前記スケジューラは、前記アクティブキューの予約バンド幅の累計値と前記リソース（バンド幅）との比（以下、予約バンド幅使用率という。）を算出する予約バンド幅使用率計算手段を備え、前記クラスタ間選択手段は、前記転送終了予定時刻を、前記各第1順位パケットの転送開始基準時刻と前記リソース占有予定時間と前記予約バンド幅使用率とを基にして算出する構成を有することを特徴とする請求項1記載のパケット転送制御装置。

【請求項3】 前記クラスタ内選択手段は、クラスタ内のパケットを、転送終了予定時刻の早い順にソートする構成を有し、前記クラスタ管理手段は、転送終了予定時刻の早い順にソートされたパケットを、各クラスタ毎に保持する構成を有することを特徴とする請求項1又は2記載のパケット転送制御装置。

【請求項4】 前記クラスタ内選択手段は、クラスタ内のパケットに対し、転送終了予定時刻の最も早い第1順位パケットのみを選出する処理と、転送終了予定時刻の早い順にソートする処理とを切り替えて行う構成を有し、

送出中のパケットの転送時間が所定時間以上要する場合は、転送終了予定時刻の早い順にソートする処理を行ない、送出中のパケットの転送時間が所定時間以下の場合は、転送終了予定時刻の最も早い第1順位パケットのみを選出する処理を行なう構成を有することを特徴とする請求項1又は2記載のパケット転送制御装置。

【請求項5】 前記クラスタ内選択手段は、クラスタ内の各パケットの前記転送開始基準時刻のみ又は前記転送開始基準時刻及び前記リソース占有予定時間を基にして、前記転送終了予定時刻の最も早いパケットを選出する構成を有することを特徴とする請求項1、2又は4記載のパケット転送制御装置。

【請求項6】 前記クラスタ内選択手段は、クラスタ内の各パケットの前記転送開始基準時刻のみ又は前記転送開始基準時刻及び前記リソース占有予定時間を基にして、前記転送終了予定時刻の早い順に、クラスタ内パケットをソートする構成を有することを特徴とする請求項3又は4記載のパケット転送制御装置。

【請求項7】 前記クラスタ分け手段は、前記リソース占有予定時間が特定されるトラフィック特性を有するアプリケーションパケットの種別を識別して、各キューの先頭パケットをクラスタに分ける構成を有することを特徴とする請求項1乃至6のいずれか1項に記載のパケット転送制御装置。

【請求項8】 前記パケット転送制御装置は、サービス品質のクラスに応じて異なるマルチキューに格納されるパケットに対して、該マルチキュー毎にそれぞれ伝送システム全体のリソースの割り当て処理が異なるスケジューラを複数備え、

それぞれのスケジューラは、各マルチキューの先頭パケットの中から転送終了予定時刻の最も早いパケットを選出する手段を備え、

更に前記マルチキュー対応のスケジューラにより選出されたパケットの中から、その転送終了予定時刻の最も早いパケットを選出するスケジューラを備えたことを特徴とする請求項1乃至7いずれか1項記載のパケット転送制御装置。

【請求項9】 前記マルチキュー対応のスケジューラのうち少なくとも一つのスケジューラは、対応するマルチキュー内の各キューに固定レートバンド幅を割り当て、転送終了予定時刻の最も早いパケットを選出する手段を備えたことを特徴とする請求項8記載のパケット転送制御装置。

【請求項10】 予め個々に定められた予約バンド幅が最小限割り当てられる複数のキューに格納されたパケットに対して、各キューの先頭パケットに、転送終了予定時刻の早い順に転送順序の重み付けを行うとともに、転送すべきパケットが格納されたアクティブキューの全てに、リソース（バンド幅）を割り当てて転送するパケット転送スケジューリング方法において、

各キューの先頭パケットの前記リソース占有予定時間に応じて、各キューの先頭パケットをクラスタに分ける過程と、

前記クラスタ内にパケットの削除又は追加があったときに、該クラスタに対して転送終了予定時刻の最も早いパケットを第1順位パケットとして選出し、保持する過程

と、クラスタ間において、前記各クラスタの第1順位パケットの中から、転送終了予定時刻の最も早いパケットを最優先転送パケットとして選出する過程とを含むことを特徴とするパケット転送スケジューリング方法。

【請求項11】 前記パケット転送スケジューリング方法において、前記アクティブキューの予約バンド幅使用率を算出する過程を含み、前記クラスタ間の選択における前記転送終了予定時刻の計算を、前記各第1順位パケットの転送開始基準時刻と前記リソース占有予定時間と前記予約バンド幅使用率とを基にして算出する過程を含むことを特徴とする請求項10記載のパケット転送スケジューリング方法。

【請求項12】 前記クラスタ内の選択において、クラスタ内のパケットを、転送終了予定時刻の早い順にソートして保持する過程を含むことを特徴とする請求項10又は11記載のパケット転送スケジューリング方法。

【請求項13】 前記クラスタ内の選択において、送出中のパケットの転送時間が所定時間以上要する場合は、転送終了予定時刻の早い順にソートする処理を行ない、送出中のパケットの転送時間が所定時間以下の場合は、転送終了予定時刻の最も早い第1順位パケットのみを選出する処理を行なう過程を含むことを特徴とする請求項10又は11記載のパケット転送スケジューリング方法。

【請求項14】 前記クラスタ内の選択において、クラスタ内の各パケットの前記転送開始基準時刻のみ又は前記転送開始基準時刻及び前記リソース占有予定時間を基にして、前記転送終了予定時刻の最も早いパケットを選出する過程を含むことを特徴とする請求項10、11又は13記載のパケット転送スケジューリング方法。

【請求項15】 前記クラスタ内の選択において、クラスタ内の各パケットの前記転送開始基準時刻のみ又は前記転送開始基準時刻及び前記リソース占有予定時間を基にして、前記転送終了予定時刻の早い順に、クラスタ内パケットをソートする過程を含むことを特徴とする請求項12又は13記載のパケット転送スケジューリング方法。

【請求項16】 前記クラスタ分けにおいて、前記リソース占有予定時間が特定されるトラフィック特性を有するアプリケーションパケットの種別を識別して、各キューの先頭パケットをクラスタに分ける過程を含むことを特徴とする請求項10乃至15のいずれか1項に記載のパケット転送スケジューリング方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、パケットを格納する複数のキューを備えたマルチキュー構成のパケット転送制御装置及びそのスケジューリング方法に関し、特に、各キューの先頭パケットの中から最も早く転送終了

することができるパケットを効率的に選出して出力するパケット転送制御装置及びそのスケジューリング方法に関する。

【0002】 パケット転送ネットワークにおいて、フロー（ユーザが転送要求する複数の連続したパケット）毎にサービス品質（QoS）を保証するとともに、伝送路等のリソース（伝送可能なバンド幅）を最大限に利用してサービスの向上を図るため、パケット転送制御装置は、フロー対応に複数のキューを備え、各キューに転送順序の重み付けを行なって、各キュー間で公平なパケット転送が行われるように制御する重み付けマルチプルキューイング、及び一時的に剰余リソース（余ったバンド幅）を各々のフロー間で公平に配分してリソースを有効利用する公平キューイング（Fair Queueing）の機能を具備することが要求されている。

【0003】 例えば、FTP（File Transfer Protocol）データ又はWWW（World Wide Web）データのようなトラヒックデータに対しては、各々のフローのデータを最大限のリソースを利用し、高速かつ公平に転送することが重要視され、ネットワークのゲートウェイのルータ等には、それらのフロー（あるいはフローの集合）毎に予約されたバンド幅を、各フローに割り当てて保証すると同時に、出力リンクにおいて余ったリソースを、それらの各フロー間で公平に再配分する制御を、高速の出力リンクに応じて高速処理することが要求されている。

【0004】 一方、VoIP（Voice over IP）のような音声トラヒックデータに対しては、リンクのバンド幅を最大かつ公平に利用するより、むしろ固定の転送レートを保証し、遅延ジッタを最小にすることが重要である。

【0005】 本発明は、このように要求されるサービス品質（QoS）が異なるパケットの転送を対象とし、1）トラヒックデータの種類毎に異なるサービス品質の要求に応じた個々のパケットの転送順序制御（スケジューリング）を高速に処理する構成と、2）これらの個々の構成を連携し、かつ全体のバンド幅を有効に利用する構成とに係り、今後の高速マルチメディアネットワークへ適用されるものである。

【0006】

【従来の技術】 公平キューイング（Fair Queueing）技術として、重み付け公平キューイング（WFQ：Weighted Fair Queueing）及び開始時間公平キューイング（SFQ：Start-time Fair Queueing）がよく知られている。

【0007】 重み付け公平キューイング（WFQ）を基本とする公平キューイング（Fair Queueing）の原理を図5を用いて説明する。同図において、5-1はパケット転送のスケジューラ、5-2は各キュー対応に備えられたメータ、5-3はそれぞれのフローに対応したN個のキュー（ $Q_1 \sim Q_N$ ）から成るマルチキューである。

【0008】 各々のキュー5-3にはフロー毎にパケッ

トが格納され、各キュー内においてはFIFOで処理されるものとし、スケジューラ5-1は各キュー5-3に格納された先頭パケットの中から、最も早く転送を終えることができるパケットを選択して該パケットを出力リンクに出力する処理を行なう。

【0009】このように重み付け公平キューイング(WFQ)による転送制御は、各キュー Q_i ($i=1, 2,$

$$F_{i,1} = \max \{ S_{i,0} + L_{i,0} / r_{i,0} ; t_{i,1} \} + L_{i,1} / r_{i,1} \quad \dots (式1)$$

【0011】ここで、 $\max \{ ; \}$ は、 $\{ \}$ 内の；で区切られた二つの値のうちの最大値を意味し、また、 $S_{i,0}$ 等は以下のとおりである。なお、下付き文字の1つ目はキュー(フロー)の番号、2つ目はキューの中での順番を意味している。

【0012】 $S_{i,0}$ は、キュー Q_i の現在の先頭パケットの一つ前のパケット(以下、直前パケットという。)の転送開始時刻、 $L_{i,0}$ は、キュー Q_i の直前パケットのパケット長、 $r_{i,0}$ は、キュー Q_i の直前パケットの転送時にキュー Q_i に割り当てられるバンド幅、 $t_{i,1}$ は、キュー Q_i の先頭パケットがキュー Q_i に到着した(格納された)時刻、 $L_{i,1}$ は、キュー Q_i の先頭パケットのパケット長、 $r_{i,1}$ は、キュー Q_i の先頭パケットの転送時にキュー Q_i に割り当てられるバンド幅である。

【0013】前述の $S_{i,0}$ 、 $L_{i,0}$ 、 $L_{i,1}$ 、 $t_{i,1}$ は、パケット転送開始時又はパケット到着時に定まる固定的な値であり、各キュー Q_i 対応のメータ5-2に記憶保持されるが、前述の $r_{i,0}$ 、 $r_{i,1}$ は、以下のように計算によって算出しなければならない。

$$r_{i,0} = (\Phi_i / \Phi_{b(0)}) \times R \quad \dots (式2) \quad 30$$

$$r_{i,1} = (\Phi_i / \Phi_{b(1)}) \times R \quad \dots (式3)$$

【0014】ここで、 Φ_i は、キュー Q_i の予約バンド幅であり、キュー Q_i のパケット転送に割り当てられるバンド幅である。 $\Phi_{b(0)}$ は、直前パケット転送時における、バックログキューの予約バンド幅の合計値、即ち $\Phi_{b(0)} = \sum \Phi_j$ ($j \in$ 直前パケット転送時のバックログキュー)であり、 $\Phi_{b(1)}$ は、先頭パケット転送時における、バックログキューの予約バンド幅の合計値、即ち $\Phi_{b(1)} = \sum \Phi_j$ ($j \in$ 先頭パケット転送時のバックログキュー)であり、 R は、リソースの全出力バンド幅であり、 $\Phi_{b(0)} \leq R$ 、 $\Phi_{b(1)} \leq R$ である。

【0015】また、前述の「バックログキュー」とは、未転送パケット(バックログ)が少なくとも1つ格納されているキューを意味し、転送動作を必要とするパケットが少なくとも1つ格納されているキューであるので「アクティブキュー」とも称される。

$$F_{0n+1,1} = \min \{ F_{i,1} \mid i \in \text{バックログキュー} \} \quad \dots (式4)$$

【0022】

【発明が解決しようとする課題】前述の(式1)の計算は、各キュー Q_i に割り当てられるバンド幅 $r_{i,0}$ 、 $r_{i,1}$

\dots 、 N) のそれぞれの先頭パケットに対して、転送終了予定時刻 $F_{i,1}$ を算出し、その値により出力順序に優先度を付け、各フロー間のパケット転送の公平性とリソースの有効利用を図っている。

【0010】各キュー Q_i のそれぞれの先頭パケットの転送終了予定時刻 $F_{i,1}$ は、次の(式1)により算出される。

【0016】各キュー Q_i に割り当てられるバンド幅 $r_{i,1}$ は、バックログキューの数によって変動し、(式2)又は(式3)から明らかのように、 $\Phi_b = R$ のときバンド幅 Φ_i となり、 $\Phi_b = \Phi_i$ のとき最大限のバンド幅 R となる。

【0017】ここで、(式1)の右辺第1項は、キュー Q_i の先頭パケットの転送開始基準時刻を表し、その直前パケットの転送終了時刻($S_{i,0} + L_{i,0} / r_{i,0}$)と先頭パケット到着時刻 $t_{i,1}$ のうち、大きい方の値を転送開始基準時刻として与える。(式1)の右辺第2項は、先頭パケットの転送に要する時間である。

【0018】そして、スケジューラ5-1は、転送終了予定時刻 $F_{i,1}$ の値の小さい順にパケットを選択して転送する制御を行う。従って、直前パケットの転送に長時間を要したキューや、転送に長時間を要する先頭パケットのキューに対しては、転送順序の優先順位が下がり、キュー間での公平なパケット転送が行われる。

【0019】このように、重み付け公平キューイング(WFQ)は、各キュー間のパケットに対して、転送終了予定時刻 $F_{i,1}$ による転送順序の重み付け制御を行って、キュー間の公平パケット転送を図るとともに、使用されていない予約バンド幅(剰余リソース)を、アクティブキューにその予約バンド幅に応じて公平に割り当ててリソースを有効利用し、アクティブキューのパケットをより高速に転送し、サービスの向上を図っている。

【0020】しかし、バックログキューの数は常時変動するため、各キュー Q_i に割り当てられるバンド幅 $r_{i,1}$ も常に変動し、各キュー Q_i の先頭パケットの転送終了予定時刻 $F_{i,1}$ も変動することとなる。

【0021】したがって、スケジューラ5-1は、一つのパケットを出力することにより、バックログキューの全ての先頭パケットに対して、前述の(式1)、(式2)及び(式3)によりそれぞれの先頭パケットの転送終了予定時刻 $F_{i,1}$ を算出し、そして、以下の(式4)に記すように、それらの中から、最小の値の転送終了予定時刻 $F_{0n+1,1}$ となるキュー Q_{0n+1} の先頭パケットを選択して転送出力する処理を行なっていた。

$t_{i,1}$ の計算を含み、それらは(式2)及び(式3)に示されるように、一つのパケットの転送のたびにバックログキューの変化により変動し、直前のパケットの出力処

理を終わった後でないと、次の先頭パケットの割り当てバンド幅 $r_{i,1}$ を正確に算出することができない。

【0023】すなわち、毎回新しい先頭パケットを出力するたびに、上記の計算をやり直さなければならず、しかも前回のパケット出力のときに計算した各キューの先頭パケット転送終了予定時刻 $F_{i,1}$ は、選択されずにキュー Q_i に残されたままとなった先頭パケットについて

$$F_{i,1} = \max \{ F_{i,0}, V(t_{i,1}) \} + L_{i,1} / \phi_{i,1} \dots (式1')$$

但し、 $V(t_{i,1}) = V(t_{i,0}) + (R / \phi_b)$

$$(t_{i,1} - t_{i,0})$$

$$V(0) = 0$$

のように解釈することができる。

【0025】バーチャルタイム $V(t)$ は、バックログキューの予約バンド幅の合計値 ϕ_b とリソースバンド幅 R とが変化するたびに再計算するが、パケットの到着時に転送終了予定時刻 $F_{i,1}$ を決定することができるので、実装が容易である。

【0026】しかし、バーチャルタイムを導入すると、転送終了予定時刻 $F_{i,1}$ の時間軸が実時間とは異なる独自のものになってしまうので、パケットの転送順序制御を行なうスケジューラを、サービス品質要求 (QoS) に応じて複数設けた場合、それぞれのスケジューラ間の時間軸が異なってしまう、スケジューラ間での最優先出力パケットの選択を行なう処理等、スケジューラ間の連携が困難となる欠点がある。

【0027】また、バーチャルタイムを採用せずに、重み付け公平キューイング (WFQ) のようなリソース公平再配分によるパケット転送制御を行なうには、転送順位決定 (スケジューリング) のための計算量がきわめて多くなってしまう、高速リンクに適用することが困難であるとしばしば指摘されている。

【0028】本発明は、フロー毎に異なるサービス品質 (QoS) を保証するとともに、伝送路のリソースを最大限に利用する公平キューイング ($Fair Queueing$) において、パケット転送のスケジューリングのための計算量を大幅に削減し、高速パケット転送処理を可能にし、しかも要求されるサービス品質 (QoS) に応じて複数設けたスケジューラ間での連携が容易な、パケット転送制御装置及びそのスケジューリング方法を提供することを目的とする。

【0029】更に具体的に詳述すると、前述の問題点に対して、転送出力する最優先のパケットを選出する計算において、キュー Q_i に割り当てられるバンド幅 $r_{i,1}$ の変化によって、直接影響を受けるパケットの範囲を限定化し、限定化したパケットのみに対して転送終了予定時刻 $F_{i,1}$ を計算し、その外のパケットに対しては過去に計算した計算結果を再利用して計算を行わず、重み付け公平キューイング (WFQ) の計算量の削減を図るものである。

【0030】

も、次のパケット出力においてその計算結果を全く流用することができず、計算量の増大を招いていた。

【0024】一方、前述の割り当てバンド幅 $r_{i,1}$ の変化による計算量増大の問題を解決するために、バーチャルタイムを導入する技術が知られている。バーチャルタイムを利用することにより、前述の(式1)は、

【課題を解決するための手段】本発明のパケット転送制御装置は、(1) 予め個々に定められた予約バンド幅が割り当てられる複数のキューに格納されたパケットに対して、各キューの先頭パケットに、転送終了予定時刻の早い順に転送順序の重み付けを行うとともに、転送すべきパケットが格納されたアクティブキューの全てに、伝送システム全体のリソース (バンド幅) を割り当てて転送制御を行うスケジューラを少なくとも備えたパケット転送制御装置において、前記スケジューラは、各キューの先頭パケットのパケット長と前記予約バンド幅との比 (以下、リソース占有予定時間という。) に応じて、各キューの先頭パケットをクラスタに分けるクラスタ分け手段と、前記クラスタ内のパケットのうちで転送終了予定時刻の最も早いパケットを第1順位パケットとして、各クラスタ毎に保持するクラスタ管理手段と、前記クラスタ内にパケットの削除又は追加があったときに、該クラスタに対して転送終了予定時刻の最も早いパケットを第1順位パケットとして選出し、前記クラスタ管理手段に保持させるクラスタ内選択手段と、前記各クラスタの第1順位パケットの中から、転送終了予定時刻の最も早いパケットを最優先転送パケットとして選出するクラスタ間選択手段とを備えたものである。

【0031】また、(2) 前記スケジューラは、前記アクティブキューの予約バンド幅の累計値と前記伝送システム全体のリソース (バンド幅) との比 (以下、予約バンド幅使用率という。) を算出する予約バンド幅使用率計算手段を備え、前記クラスタ間選択手段は、前記転送終了予定時刻を、前記各第1順位パケットの転送開始基準時刻と前記リソース占有予定時間と前記予約バンド幅使用率とを基にして算出する構成を有するものである。

【0032】また、(3) 前記クラスタ内選択手段は、クラスタ内のパケットを、転送終了予定時刻の早い順にソートする構成を有し、前記クラスタ管理手段は、転送終了予定時刻の早い順にソートされたパケットを、各クラスタ毎に保持する構成を有するものである。

【0033】また、(4) 前記クラスタ内選択手段は、クラスタ内のパケットに対し、転送終了予定時刻の最も早い第1順位パケットのみを選出する処理と、転送終了予定時刻の早い順にソートする処理とを切り替えて行う構成を有し、送出中のパケットの転送時間が所定時間以上要する場合は、転送終了予定時刻の早い順にソートする処理を行ない、送出中のパケットの転送時間が所定時

間以下の場合、転送終了予定時刻の最も早い第1順位パケットのみを選出する処理を行なう構成を有するものである。

【0034】また、(5)前記クラスタ内選択手段は、クラスタ内の各パケットの前記転送開始基準時刻のみ又は前記転送開始基準時刻及び前記リソース占有予定時間を基にして、前記転送終了予定時刻の最も早いパケットを選出する構成を有するものである。

【0035】また、(6)前記クラスタ内選択手段は、クラスタ内の各パケットの前記転送開始基準時刻のみ又は前記転送開始基準時刻及び前記リソース占有予定時間を基にして、前記転送終了予定時刻の早い順に、クラスタ内パケットをソートする構成を有するものである。

【0036】また、(7)前記クラスタ分け手段は、前記リソース占有予定時間が特定されるトラフィック特性を有するアプリケーションパケットの種別を識別して、各キューの先頭パケットをクラスタに分ける構成を有するものである。

【0037】また、(8)前記パケット転送制御装置は、サービス品質のクラスに応じて異なるマルチキューに格納されるパケットに対して、該マルチキュー毎にそれぞれ伝送システム全体のリソースの割り当て処理が異なるスケジューラを複数備え、それぞれのスケジューラは、各マルチキューの先頭パケットの中から転送終了予定時刻の最も早いパケットを選出する手段を備え、更に前記マルチキュー対応のスケジューラにより選出されたパケットの中から、その転送終了予定時刻の最も早いパケットを選出するスケジューラを備えたものである。

【0038】また、(9)前記マルチキュー対応のスケジューラのうち少なくとも一つのスケジューラは、対応するマルチキュー内の各キューに固定レートのバンド幅を割り当て、転送終了予定時刻の最も早いパケットを選出する手段を備えたものである。

【0039】また、本発明のパケット転送スケジューリング方法は、(10)予め個々に定められた予約バンド幅が最小限割り当てられる複数のキューに格納されたパケットに対して、各キューの先頭パケットに、転送終了予定時刻の早い順に転送順序の重み付けを行うとともに、転送すべきパケットが格納されたアクティブキューの全てに、伝送システム全体のリソース(バンド幅)を割り当てて転送するパケット転送スケジューリング方法において、各キューの先頭パケットの前記リソース占有予定時間に応じて、各キューの先頭パケットをクラスタに分ける過程と、前記クラスタ内のパケットに削除又は追加があったときに、該クラスタに対して転送終了予定時刻の最も早いパケットを第1順位パケットとして選出し、保持する過程と、クラスタ間において、前記各クラスタの第1順位パケットの中から、転送終了予定時刻の最も早いパケットを最優先転送パケットとして選出する過程とを含むものである。

【0040】また、(11)前記パケット転送スケジューリング方法において、前記アクティブキューの予約バンド幅使用率を算出する過程を含み、前記クラスタ間の選択における前記転送終了予定時刻の計算を、前記各第1順位パケットの転送開始基準時刻と前記リソース占有予定時間と前記予約バンド幅使用率とを基にして算出する過程を含むものである。

【0041】また、(12)前記クラスタ内の選択において、クラスタ内のパケットを、転送終了予定時刻の早い順にソートして保持する過程を含むものである。

【0042】また、(13)前記クラスタ内の選択において、送出中のパケットの転送時間が所定時間以上要する場合は、転送終了予定時刻の早い順にソートする処理を行ない、送出中のパケットの転送時間が所定時間以下の場合は、転送終了予定時刻の最も早い第1順位パケットのみを選出する処理を行なう過程を含むものである。

【0043】また、(14)前記クラスタ内の選択において、クラスタ内の各パケットの前記転送開始基準時刻のみ又は前記転送開始基準時刻及び前記リソース占有予定時間を基にして、前記転送終了予定時刻の最も早いパケットを選出する過程を含むものである。

【0044】また、(15)前記クラスタ内の選択において、クラスタ内の各パケットの前記転送開始基準時刻のみ又は前記転送開始基準時刻及び前記リソース占有予定時間を基にして、前記転送終了予定時刻の早い順に、クラスタ内パケットをソートする過程を含むものである。

【0045】また、(16)前記クラスタ分けにおいて、前記リソース占有予定時間が特定されるトラフィック特性を有するアプリケーションパケットの種別を識別して、各キューの先頭パケットをクラスタに分ける過程を含むものである。

【0046】

【発明の実施の形態】図1は本発明の重み付け公平キューイング(WFQ)によるパケット転送制御の原理説明図である。同図において、1-1はパケットをクラスタ化して転送制御するスケジューラ、1-2は各フロー対応にパケットを格納するN個のキュー($Q_1 \sim Q_N$)から成るマルチキューである。

【0047】各キュー Q_i ($i=1, 2, \dots, N$)の内部構成は、本発明のパケット転送制御に直接影響がないので、各キュー Q_i は、一般的なFIFO(First In First Out)によるパケット格納処理を行う構成を有するものと仮定する。

【0048】スケジューラ1-1は、各キュー Q_i の先頭パケット(図の網掛け部分A)に対して、各パケットの転送に要するリソース占有予定時間 $\beta_{i,1}$ をもとに、クラスタ分けを行なう。同図は、各キュー Q_i の先頭パケットをM個のクラスタ $C_1 \sim C_M$ に分けて転送制御する例を示している。

【0049】各パケット転送に要するリソース占有予定時間 $\beta_{i,1}$ は、各キュー Q_i の先頭パケットのパケット長 $L_{i,1}$ と、各キュー Q_i の予約バンド幅 Φ_i とから、以下の式

$$\beta_{i,1} = L_{i,1} / \Phi_i$$

により算出されるもので、キュー Q_i に最低限割り当てられる予約バンド幅 Φ_i で、パケット長 $L_{i,1}$ のパケットを転送するのに要する時間である。

【0050】実際には、前述したように重み付け公平キューイング(WFQ)により、不使用の予約バンド幅があれば、そのバンド幅も所定の比率でアクティブキューに割り当てるため、リソース占有時間は上記の $\beta_{i,1}$ より短縮されることとなるので、上記の $\beta_{i,1}$ をリソース占有予定時間と称する。

【0051】スケジューラ1-1は、各先頭パケットのリソース占有予定時間 $\beta_{i,1}$ の値に応じて、各先頭パケットをそれぞれM個のクラスタ $C_1 \sim C_M$ にクラスタ分けする。同図は、クラスタ C_1 にキュー Q_3 の先頭パケット $P_{3,1}$ とキュー Q_3 の先頭パケット $P_{3,1}$ とがクラスタ化され、クラスタ C_2 にキュー Q_2 の先頭パケット $P_{2,1}$ とキュー Q_1 の先頭パケット $P_{1,1}$ とがクラスタ化され、クラスタ C_M にキュー Q_1 の先頭パケット $P_{1,1}$ がクラスタ化された場合の例を示している。

【0052】スケジューラ1-1は、クラスタ内に新たにキューの先頭パケットを組み入れる際、そのクラスタ内で最も小さな値の転送終了予定時刻 $F_{i,1}$ を持つ先頭パケットを、このクラスタ内で最優先に出力すべき第1順位パケットとして登録する。

【0053】また、クラスタ内の第1順位パケットが転

$$\alpha_{i,1} = \max \{ S_{i,0} + L_{i,0} / r_{i,0} ; t_{i,1} \} \quad \dots (式5)$$

【0058】ここで、 $S_{i,0} + L_{i,0} / r_{i,0}$ は、先頭パケットの一つ前のパケット(直前パケット)の転送終了予定時刻、 $t_{i,1}$ は当該先頭パケットの到着時刻である。当該先頭パケットの転送開始基準時刻 $\alpha_{i,1}$ は、直前パケットの転送終了予定時刻と当該先頭パケットの到着時刻の値のうち、大きい方の値となる。

【0059】リソース占有予定時間 $\beta_{i,1}$ は前述したとおり、

$$\beta_{i,1} = L_{i,1} / \Phi_i \quad \dots (式6)$$

により計算される。リソース占有予定時間 $\beta_{i,1}$ は、キュー Q_i のパケットを、その最低限の割り当てバンド幅である予約バンド幅 Φ_i で転送とした場合、転送に要するリソース占有時間である。キュー Q_i に割り当てられるバンド幅は、バックログキューの変化により変動するため、該リソース占有時間も変動する。したがって、リソース占有予定時間 $\beta_{i,1}$ は、実際のリソース占有の予定時間ではなく、以下に述べるように、転送終了予定時刻 $F_{i,1}$ の計算に用いられる基準値の一つである。

【0060】スケジューラ1-1は、更に先頭パケット

送パケットとして選出され、出力リンクへ送出された際も、クラスタ内の残りのパケットの中から最も小さな値の転送終了予定時刻 $F_{i,1}$ を持つパケットを、クラスタ内の第1順位パケットとして登録する。

【0054】図1では、クラスタ C_2 からパケット $P_{2,1}$ が送出され、キュー Q_2 から次のパケット $P_{2,2}$ が、先頭パケットとして同じクラスタ C_2 に組み込まれた例を示しているが、このように一つのクラスタを構成するパケットが変動したとき、そのクラスタ内のパケット間(図の網掛け部分B)で、転送終了予定時刻 $F_{i,1}$ の値の最も小さいパケットを、第1順位パケットとして登録する。

【0055】そして、スケジューラ1-1は、それぞれのクラスタの第1順位パケット(図の例においては網掛け部分Cのパケット $P_{3,1}$ 、 $P_{2,1}$ 、 \dots 、 $P_{1,1}$)に対して、転送終了予定時刻 $F_{i,1}$ を計算し、その値が最も小さいパケットを最優先に転送するパケットとして選出し、該パケットを出力リンクから送出する制御を行なう。

【0056】本発明のパケット転送制御の原理について更に詳しく説明する。図1に示したスケジューラ1-1は、各キュー Q_i に対応したメータ(図示省略)を備え、その各メータには、キューの先頭パケットの転送終了予定時刻 $F_{i,1}$ を計算するために必要なパラメータが設定される。

【0057】キュー Q_i のメータに設定されるパラメータは、先頭パケットの転送開始基準時刻 $\alpha_{i,1}$ とリソース占有予定時間 $\beta_{i,1}$ である。先頭パケットの転送開始基準時刻 $\alpha_{i,1}$ は、以下の式により算出される。

の転送終了予定時刻 $F_{i,1}$ を計算するためのパラメータであるバックログキューによる予約バンド幅使用率 γ を計算する。予約バンド幅使用率 γ は、リソース(全出力リンクのバンド幅) R に対する、各バックログキュー(アクティブキュー)により使用されるそれぞれの予約バンド幅の合計値 $\Phi_{b(i)}$ の比率であり、以下の式により算出される。

$$\gamma = \Phi_{b(i)} / R \quad \dots (式7)$$

【0061】前述した(式1)の先頭パケットの転送終了予定時刻 $F_{i,1}$ は、(式3)、(式5)、(式6)及び(式7)により、パラメータ $\alpha_{i,1}$ 、 $\beta_{i,1}$ 及び γ を用いた以下の(式8)のように変形することができる。
$$F_{i,1} = \alpha_{i,1} + L_{i,1} / \{ (\Phi_i / \Phi_{b(i)}) \times R \} = \alpha_{i,1} + (L_{i,1} / \Phi_i) \times (\Phi_{b(i)} / R) = \alpha_{i,1} + \beta_{i,1} \times \gamma \quad \dots (式8)$$

【0062】ここで、転送開始基準時刻 $\alpha_{i,1}$ とリソース占有予定時間 $\beta_{i,1}$ は、該パケットがキューの先頭になった時点で算定され、以降、変化することのない固定値であるが、予約バンド幅使用率 γ は、バックログキューの変化により変動する値である。

【0063】上記の(式8)による転送終了予定時刻 $F_{i,1}$ と予約バンド幅使用率 γ との関係をグラフに表すと、図2の(A)の斜線で示すように、転送終了予定時刻 $F_{i,1}$ は γ の変動に対して直線的に変動する。

【0064】図2の(A)は、キュー Q_i ($i=1\sim 4$)の各先頭パケットの転送終了予定時刻 $F_{i,1}$ ($i=1\sim 4$)を例示的に示している。図の斜線と縦軸の交点が $\alpha_{i,1}$ であり、斜線の勾配が $\beta_{i,1}$ である。

【0065】図2の(A)から分かるように、斜線の勾配 $\beta_{i,1}$ に大きな差がないパケット同士の転送終了予定時刻 $F_{i,1}$ は、予約バンド幅使用率 γ が変化しても、それらの転送終了予定時刻 $F_{i,1}$ の順序が入れ替わる確率が低く、仮に入れ替わってもその時刻差は微小なものである。

【0066】図2(A)において、 $i=1$ と $i=3$ の $F_{i,1}$ の勾配 $\beta_{i,1}$ に大差がなく、また $i=2$ と $i=4$ の $F_{i,1}$ の勾配 $\beta_{i,1}$ に大差がない。このような場合、予約バンド幅使用率 γ の変化に対し、 $i=2$ と $i=4$ の $F_{i,1}$ の大小関係は変化せず、また、 $i=1$ と $i=3$ の $F_{i,1}$ の大小関係は、途中で交差して入れ替わっているが、勾配 $\beta_{i,1}$ が近似しているため、 $i=1$ と $i=3$ の $F_{i,1}$ の値自体に、大きな差異は生じない。

【0067】図2の(B)は勾配 $\beta_{i,1}$ のみを表した図である。この勾配 $\beta_{i,1}$ に関して幾つか(例えばM個)レンジを予め定め、(式6)によって計算した勾配 $\beta_{i,1}$ がどのレンジに属するかによって、前述したように各キューの先頭パケットに対してクラスタリングを行なう。

【0068】このように先頭パケットのクラスタリングを行なうと、同一クラスタ内に属するパケットの転送終了予定時刻 $F_{i,1}$ の順序に関して、 γ の変化に対する依存度を微小なものとすることができる。

【0069】例えば、図2において、 $i=1$ と $i=3$ のキューの先頭パケットを同一クラスタとし、また、 $i=2$ と $i=4$ のキューの先頭パケットを同一クラスタとしてクラスタリングを行なうと、 $\gamma=\gamma_1$ のときに計算した各キュー Q_i のパケットの転送終了予定時刻 $F_{i,1}$ は、 $\gamma=\gamma_2$ となった時点でも、各クラスタ内のそれぞれのパケットの転送終了予定時刻 $F_{i,1}$ の順序が変化する確率は低いし、たとえ変化があってもその時刻差は小さいので、転送順序に誤差を生じたとしてもその影響は小さい。

【0070】毎回のパケット転送出力時のパケット選択処理は、

①直前にパケットを出力したクラスタ又は新たな先頭パケットが追加されたクラスタ(図1の網掛け部分B)内の全てのパケットに対し、前述の(式8)による転送終了予定時刻 $F_{i,1}$ を計算して転送終了予定時刻 $F_{i,1}$ が最小となる第1順位パケットを選出する処理(クラスタ内選択処理)と、

②各クラスタの第1順位パケット間(図1の網掛け部分C)に対して、同様に(式8)による転送終了予定時刻 $F_{i,1}$ を計算し、前述の(式4)によりその値が最小となるパケットを選出する処理(クラスタ間選択処理)とを行う。

【0071】そして、クラスタ間選択処理により選出されたパケットを、最優先転送パケットとして出力リンクへ送出する処理を行うだけで済むので、従来、全てのバックログキューの先頭パケットに対して(式8)の計算を行ったのに比し、計算量を大幅に削減することができる。

【0072】例えば、バックログキューの数が100個あったとすると、それらの先頭パケットは100個あり、従来は、100個のパケットに対して(式8)による転送終了予定時刻 $F_{i,1}$ を計算して、その最小値のパケットを選出する処理を行っていたのに対し、本発明では、100個の先頭パケットが10個のクラスタにそれぞれ平均して10個ずつクラスタ化されたとすると、平均10個のパケットのクラスタ内選択処理と、10個のクラスタについてのクラスタ間選択処理とを行えばよく、すなわち、20個のパケットに対してのみ(式8)による転送終了予定時刻 $F_{i,1}$ を計算すればよく、計算量を大幅に削減できるとともに、その最小なものを選択する処理においても、選択対象の個数が大幅に減少化されているので、選択処理における処理量も低減化することができる。

【0073】一般に、キューの数をNとし、クラスタの数をMとすると、クラスタ内のパケットの数は平均で N/M となる。すなわち、本発明のようにクラスタ分けを行なうことにより、従来N個のパケットに対して行っていた計算を、 N/M 個のパケットに対するクラスタ内処理と、M個のクラスタに対するクラスタ間処理のみで済み、 $(N/M+M)$ 個のパケットに対して計算を行なえばよく、計算量を大きく削減することができる。

【0074】なお、前述のクラスタ内選択処理において、(式8)による $\alpha_{i,1}$ 、 $\beta_{i,1}$ 及び γ を用いた転送終了予定時刻 $F_{i,1}$ の計算によって第1順位パケットを選出する処理に替えて、クラスタ内では各パケットのリソース占有予定時間 $\beta_{i,1}$ が近似した値であり、予約バンド幅使用率 γ に対する依存性は低いことから、転送開始基準時刻 $\alpha_{i,1}$ の値を第1ソートキーとして小さい順にパケットをソートし、次に、同一又は近似した値の転送開始基準時刻 $\alpha_{i,1}$ を持つパケットに対して、リソース占有予定時間 $\beta_{i,1}$ を第2ソートキーとして小さい順にソートすることにより、予約バンド幅使用率 γ を使用することなく、最小値付近の転送終了予定時刻 $F_{i,1}$ のパケットを簡易な処理で選出することができる。

【0075】更に、転送開始基準時刻 $\alpha_{i,1}$ の値のみを用い、その値の最も小さいパケットを第1順位パケットとして選出しても、選出結果に大きな誤差を生じること

はなく、最小値付近の転送終了予定時刻 $F_{i,1}$ のバケットを、更に簡易な処理で選出することができる。この場合は、 $\beta \times \gamma$ の計算を行なわないので、クラスタ内の計算量を大幅に削減することができる。

【0076】また、クラスタ内選択処理において、前述の $\alpha_{i,1}$ 、 $\beta_{i,1}$ 及び γ を用いた選出又は $\alpha_{i,1}$ 及び $\beta_{i,1}$ のみを用いた選出を行った際、選出された第1順位バケットのみを登録しておき、クラスタに新たなバケットが追加されたときに、登録済みの第1順位バケットと新たなバケットが追加されたバケットとの転送終了予定時刻 $F_{i,1}$ に関する大小比較を行って新たな第1順位バケットを決定して登録する手法と、クラスタ内のバケット全てに対して転送終了予定時刻 $F_{i,1}$ の小さい順にソートしておく手法と、更にそれらの手法をクラスタ内選択処理に許容される処理時間に応じて、切り替える手法を適用することができる。

【0077】一般に、バケットの長さにはばらつきがあり、この特性を利用し、長いバケットなど転送に多くの時間を要するバケットを送出しているとき、次の出力バケットを選出する計算を行う時間も多く与えられるので、その間に各クラスタ内のバケットをソートしておき、短いバケットを高速に送出するときや、削除のみで追加バケットがないクラスタにおける選択処理では、過去にソートした結果をそのまま流用するか、又は第1順位バケットだけを選ぶ処理に切り替え、処理量を低減し、連続する短いバケットを出力する場合でも、高速な選択処理を、精度の劣化なく又は僅かな精度の低下を伴うだけで行うことができる。

【0078】なお、ソート処理を開始するに先立って、転送に長時間を要するバケットかどうかの判定は、バケット長 L と全出力バンド幅 R との割算 L/R を計算し、その値と閾値との比較により判定することができ、またバケット長のみと閾値との比較により簡易な判定を行ってもよい。

【0079】更に、クラスタ内選択処理において、第1順位バケットを選出する際に、その転送終了予定時刻 $F_{i,1}$ を算出する場合は、クラスタ間選択処理においては、クラスタ内選択処理を行ったクラスタを除いて、その他のクラスタの第1順位バケットに対してのみ転送終了予定時刻 $F_{i,1}$ を計算し、それらの値のうちの最小なものとはクラスタ内選択処理において算出した転送終了予定時刻 $F_{i,1}$ とを比較して、転送すべき最優先のバケットを選出することができる。

【0080】更に、音声やビデオ情報のようなマルチメディアのアプリケーションバケットの転送においては、一般的に、その品質要求に応じて、バケットを幾つかのグレードに分けて転送する。

【0081】そして、各グレードでは予約バンド幅 ϕ_i とバケット長 L_i といったトラフィック特性が均一であり、アプリケーションの種類を識別することによって、

リンク占有予定時間 β_i ($\beta_i = L_i / \phi_i$)が一義的に定まる場合がある。このような場合は、アプリケーションバケットのトラフィック特性種別を識別してバケットのクラスタ化を簡易かつ正確に行うことができる。

【0082】次に、本発明の重み付け公平キューイング(WFQ)によるバケット転送制御装置の具体的な構成について説明する。ここで、バケット転送制御装置は、フロー毎にキューイングを行なうものとし、各キューの内部ではFIFOによりバケットの格納処理を行い、更に、バケットデータの本体は、図示を省略したバッファに保存され、以下に説明するバケット転送制御装置は、バケットデータ本体へのポインタと転送制御に関連するパラメータだけを処理の対象とするものとする。

【0083】図3は本発明のバケット転送制御装置の構成を示す図である。同図において、3-1はマルチキュー管理手段、3-2は予約バンド幅使用率 γ 計算手段、3-3はクラスタ分け手段、3-4はクラスタ管理手段、3-5はクラスタ内選択手段、3-6はクラスタ間選択手段、3-7は出力制御手段である。

【0084】マルチキュー管理手段3-1は、フロー対応のキュー Q_i ($i=1 \sim N$)と各キュー Q_i 対応のメータとを備える。各キュー Q_i には、受信したバケットのフロー番号FlowNo. (j)とバケット長PacketLen. (L)と到着時間(t)とバケットデータ本体を記憶するバッファのアドレスのポインタPointerとから成るバケットタグを到着順に格納する。

【0085】各キュー Q_i 対応のメータには、そのキュー Q_i の先頭バケットの転送開始基準時刻 $\alpha_{i,1}$ とそのキュー Q_i が予約しているバンド幅 ϕ_i とを保存する。マルチキュー管理手段3-1は、未転送バケット(バックログ)が少なくとも1つ格納されているアクティブ状態のキュー Q_i と各キュー対応のメータとを管理する機能を備え、出力要求によって、対応するキューの先頭バケットのタグ及びそのメータに記憶されているパラメータを出力する機能を備える。

【0086】更に、キューのアクティブ状態が変化した時、変化した予約バンド幅($\pm \phi$)を出力する機能を備える。例えば、キュー Q_i が非アクティブ状態からアクティブ状態へ変化したときは、その予約バンド幅 ϕ_i の増加分($+\phi_i$)、逆にアクティブ状態から非アクティブ状態へ変化したときは、その予約バンド幅 ϕ_i の減少分($-\phi_i$)を、予約バンド幅使用率 γ 計算手段3-2に出力する。

【0087】予約バンド幅使用率 γ 計算手段3-2は、 ϕ_0 計算部と γ 計算部とを備える。 ϕ_0 計算部はマルチキュー管理手段3-1から、変化した予約バンド幅($\pm \phi$)を入力し、アクティブキュー(バックログキュー)の予約バンド幅の累計値 ϕ_0 ($\phi_0 = \phi_0 \pm \phi$)を計算する。

【0088】 γ 計算部は、アクティブキュー(バックロ

グキュー)の予約バンド幅の累計値 Φ_i を、リソース(出力リンク)全体のバンド幅 R で除算し、アクティブキュー(バックログキュー)全体により使用される予約バンド幅 Φ_i の、バンド幅 R に対する比率 γ を計算する。

【0089】クラスタ分け手段3-3は、 β_i 計算部とクラスタ分け処理部とを備える。 β_i 計算部は、キュー Q_i の先頭のパケットのパケット長 L_i をキュー Q_i の予約バンド幅 Φ_i で除算し(L_i/Φ_i)、該先頭パケットのリソース占有予定時間 β_i を計算する。

【0090】クラスタ分け処理部は、当該パケットのリソース占有予定時間 β_i の値に基づき、その値の属するレンジに対応して当該パケットをクラスタ $C_1 \sim C_M$ に振り分ける処理を行う。

【0091】クラスタ管理手段3-4は、クラスタ内選択手段3-5と連携し、各クラスタ C_j ($j=1 \sim M$)において、クラスタ内の各パケットの転送終了予定時刻 F_i のうち、その値の一番小さいパケットがそのクラスタの第1順位パケットとして選択されている状態を維持する。

【0092】また、クラスタ内のパケットを、転送終了予定時刻 F_i の値の小さい順にソートされた状態に維持するための、次の順位のパケットのポインタを示すパラメータNCP(Next Cluster Pointer)をクラスタ内のパケットに付加し、該パラメータNCPにより、クラスタ内においてパケットをソート順にリンクさせておくことができる。

【0093】クラスタ内選択手段3-5は、クラスタ管理手段3-4に保管されているクラスタの一つに対して、クラスタ管理手段3-4からの要求により、該クラスタ内の各パケットの転送終了予定時刻 F_i ($=\alpha_i + \beta_i \times \gamma$)を計算し、各パケットの転送終了予定時刻 F_i のうち、その値の一番小さいパケットを第1順位パケットとして選出し、クラスタ管理手段3-4に返送する機能を有する。また、その第1順位パケットの転送終了予定時刻 F_i を出力制御手段3-7に出力する機能を有する。

【0094】クラスタ内選択手段3-5は、前述したように α_i と β_i とから選出する手法や、ソートする手法、又はそれらを処理時間に応じて適宜切り替える手法を適用することができる。

【0095】クラスタ間選択手段3-6は、前述のクラスタ $C_1 \sim C_M$ (ただし、直前に出力したクラスタを除く。)のそれぞれの第1順位パケットの中から、転送終了予定時刻 F_i ($=\alpha_i + \beta_i \times \gamma$)の一番小さい値のパケットを計算して選出し、その転送終了予定時刻 F_i を出力制御手段3-7に出力する。

【0096】出力制御手段3-7は、クラスタ内選択手段3-5及びクラスタ間選択手段3-6から入力される転送終了予定時刻 F_i を比較し、その値の小さいパケッ

トを転送すべき最優先パケットとして選出し、そのパケットタグを出力する。

【0097】そして、転送されることとなったパケットのタグが格納されていたキュー Q_i に対して、次のパケットを先頭パケットとして同様の処理を行なうために、出力制御手段3-7は、マルチキュー管理手段3-1、予約バンド幅使用率 γ 計算手段3-2及びクラスタ分け手段に対して、次の先頭パケットに対する処理を起動する。

10 【0098】以上、本発明のパケット転送制御装置の各機能手段について説明したが、その処理動作について、パケット出力処理とパケット入力処理とに分けて、以下に同じ図3を参照して説明する。なお図中、○印内の番号は、以下の説明における処理動作に付した()内の番号に対応し、当該処理動作に関連する情報が入出力される機能モジュール間のリンク位置を表している。更に、図中、各リンク又は機能モジュールで入出力又は保持される情報の内容を、a)~h)の枠内に引き出し線を付して記している。

20 【0099】[1]パケット出力処理について
出力制御手段3-7によるパケット出力終了(10)と次のパケットの出力許可(11)の条件が満足されると、出力制御手段3-7から、マルチキュー管理手段3-1、予約バンド幅使用率 γ 計算手段3-2及びクラスタ分け手段3-3に、直前に出力したパケットが属していたキュー Q_i の次のパケットを先頭パケットとして処理する要求(12)を送出する。

30 【0100】マルチキュー管理手段3-1は、キュー Q_i に未転送パケットが格納されていれば、その新たな先頭パケットを呼び出し、該先頭パケットの転送開始基準時刻 α_i を計算し、メータに保存する。ただし、もし当該キュー Q_i が空となっていれば、マルチキュー管理手段3-1はなにも出力しない。

【0101】転送開始基準時刻 α_i は、このパケットの到着時刻 t と直前に送出されたパケットの転送終了予定時刻 F_{i-1} とを比較し、 $t \geq F_{i-1}$ であれば、 $\alpha_i = t$ として設定され、 $t < F_{i-1}$ であれば、 $\alpha_i = F_{i-1}$ として設定される。

40 【0102】マルチキュー管理手段3-1は、パケット長 L_i 及び予約バンド幅 Φ_i のパラメータをクラスタ分け手段3-3に出力する(15)。また、マルチキュー管理手段1-3は、キューのアクティブ状態変化によって引き起こされた使用予約バンド幅の変化($\pm \Phi$)を、予約バンド幅使用率 γ 計算手段3-2に出力する(14)。

50 【0103】クラスタ分け手段3-3は、パラメータが入力されると、リソース占有予定時間 β を計算し(2)、この値に応じて新たな先頭パケットのクラスタ分け処理(3)を行ない、クラスタ管理手段3-4に出力する。

【0104】予約バンド幅使用率 γ 計算手段3-2は、バックログキューの予約バンド幅の累計値 Φ_i を計算し(13)、出力制御手段から3-7から出力要求(12)があるとき、予約バンド幅の累計値 Φ_i を γ 計算部へ出力し、 γ 計算部は、予約バンド幅の累計値 Φ_i とリソース全体のバンド幅 R との比率である予約バンド幅使用率 γ を計算し(4)、クラスタ内選択手段3-5及びクラスタ間選択手段3-6に出力する。

【0105】クラスタ管理手段3-4は、最優先出力パケットとして選出されたパケットが属していたクラスタ C_j が出力制御手段3-7から通知されると、次の先頭パケットのクラスタ分け処理が完了する期間だけ待って、そのクラスタ C_j に属する全てのパケットを、クラスタ内選択手段3-5に送出し(7)、クラスタ内選択手段3-5は、それらのパケットの中から、新たに第1順位パケットを選出してクラスタ管理手段に返送する(9)。

【0106】また、次の先頭パケットが他のクラスタ C_k にクラスタ化された場合には、他のクラスタ C_k に対しても同様に、クラスタ内選択手段3-5に送出し(7)、クラスタ内選択手段3-5は、それらのパケットの中から、新たに第1順位パケットを選出してクラスタ管理手段に返送する(9)。

【0107】クラスタ管理手段3-4は、クラスタ内選択手段3-5に最後に送出したクラスタ C_j 以外の各クラスタの第1順位パケットのパラメータ(α_i 、 β_i)を、クラスタ間選択手段3-6に出力する(5)。

【0108】クラスタ内選択手段3-5は、クラスタ管理手段3-4から入力(7)したパケットのパラメータ α_i 、 β_i 及び新たな予約バンド幅使用率 γ によって、各パケットの転送終了予定時刻 F_i (= $\alpha_i + \beta_i \times \gamma$)を計算して、その中から転送終了予定時刻 F_i が最小となるパケットを第1順位パケットとして選出し、出力制御手段3-7に出力する(8)とともに、クラスタ管理手段にも返送する(9)。

【0109】クラスタ間選択手段3-6は、入力された各クラスタの第1順位のパケットのパラメータ(α_i 、 β_i)及び新たな予約バンド幅使用率 γ によって、転送終了予定時刻 F_i (= $\alpha_i + \beta_i \times \gamma$)を計算し、その中から最小の F_i のパケットを最優先候補パケットとして選出し、この最優先候補パケットを出力制御手段3-7に出力する(6)。

【0110】出力制御手段3-7は、クラスタ間選択手段3-6及びクラスタ内選択手段3-5から入力(6)(8)された転送終了予定時刻 F_i のうち、一番小さい値のパケットを最優先転送パケットとして選択して出力する(10)と同時に、次の出力パケットに対する処理要求を行なう(12)。

【0111】〔2〕パケット入力処理

マルチキュー管理手段3-1は、転送サービスを行なう

パケットのパケットタグを受信し(1)、該パケットタグをフロー対応のキュー Q_i に格納する。このとき、該キュー Q_i が、アクティブ状態のキューであった場合は、既に先頭パケットが存在しているので、新たに受信したパケットのタグを該キュー Q_i に追加するだけで入力処理を終了する。

【0112】空の状態のキュー Q_i にパケットが到着したときは、このパケットが先頭パケットとなるので、マルチキュー管理手段3-1は、該キュー Q_i のメータにパケットの転送開始基準時刻 α_i を計算して設定し、そのパケット長 L_i 及びキュー Q_i の予約バンド幅 Φ_i をクラスタ分け手段3-3に出力し(15)し、使用予約バンド幅の変動分(増加分 Φ_i)を、予約バンド使用率 γ 計算手段3-2に出力する。

【0113】クラスタ分け手段3-3は、 β 計算部で L_i / Φ_i によりリソース占有予定時間 β_i を計算し、クラスタ分け処理部でリソース占有予定時間 β_i の値に従ってパケットのクラスタ分けを行い、クラスタ管理手段3-4に出力する(3)。

【0114】クラスタ管理手段3-4では、新たにパケットが追加されたクラスタ C_j に対して、クラスタ内選択手段3-5と連携して該クラスタ C_j の第1順位パケットと、新たに追加されたパケットとのパラメータ(α 、 β)と予約バンド使用率 γ とを基に、転送終了予定時刻 F_i を計算し、転送終了予定時刻 F_i の最も小さいパケットを新たな第1順位パケットとして選択して保持する。

【0115】図4は、本発明における複数のスケジューラによるパケット転送制御の説明図である。同図において、4-1は入力リンク、4-2はルータ(クラシファイヤ)、4-3は第1のマルチキュー、4-4は第2のマルチキュー、4-5は第1のサブスケジューラ、4-6は第2のサブスケジューラ、4-7はその上位のメインスケジューラである。

【0116】図4は、低ジッタ(Low jitter)であることが重要視される音声パケット等の転送と、低遅延(Low Delay)及び最大転送能力が重要視されるWWWやFTP等のデータパケットの転送のように、異なるサービス品質(QoS)が要求されるアプリケーションのトラヒックパケットが混在した場合のパケットスケジューリングについての説明図である。

【0117】入力リンク4-1から入力される音声パケット(Voice)、画像パケット(Video)、データパケット(Data)等、サービス品質(QoS)クラスの異なるトラヒックデータが入力され、ルータ4-2は、入力されたパケットのヘッダ情報を解析することによって、そのパケットが属するフローに分類し、対応するキューに格納する。

【0118】各キューは、対応する各フローの要求するサービス品質に応じて、それぞれ異なるスケジューラに

よって制御される複数のマルチキューとして構成される。その第1のマルチキュー4-3は、音声パケット(Voice)、画像パケット(Video)のように、低ジッタが要求されるパケットを格納し、第1のサブスケジューラ(S1)4-5によって、所定の出力レートが保証されるようにスケジューリング制御される。

【0119】第1のサブスケジューラ(S1)4-5は、キュー間で余ったバンド幅を再配分する必要はないので、前述の(式1)において、 $r_{i,n} = r_{i,1} = \phi_i$ 10 として固定し、各キューの先頭パケットの転送終了予定時刻 $F_{i,1}$ を計算することにより、その値の最も小さいパケットを最優先出力パケットとして選出する。

【0120】第1のサブスケジューラ(S1)4-5は、割り当てバンド幅 $r_{i,1}$ が固定値であるため、計算処理の負担が軽く、音声パケットのような短いパケットで構成されるフローの処理に対しても、高速に追従して優先選択処理を行なうことができる。

【0121】一方、データパケットのように、低遅延が要求されるパケットは、第2のマルチキュー4-4に格納され、第2のサブスケジューラ(S1)4-6によっ 20 て、重み付け公平キューイング(WFQ)により余剰バンド幅が公平に割り当てられスケジューリング制御される。

【0122】重み付け公平キューイング(WFQ)は、前述したようにパケットの最小遅延保証及びフロー間での余剰バンド幅の公平再分配において優れた特長を持っている。第2のサブスケジューラ(S1)4-6は、各キューの先頭パケットの転送終了予定時刻 $F_{i,1}$ の計算において、前述したようにクラスタ化を行い、計算対象 30 の先頭パケットを限定化して高速に計算処理を行い、その値の最も小さいパケットを最優先出力パケットとして選出する。

【0123】上位のメインスケジューラ(S)4-7は、第1及び第2のサブスケジューラ(S1, S2)間の連携を行い、その両者から選出されたパケットの出力候補に対して、さらにその転送終了予定時刻 $F_{i,1}$ の最も小さいパケットを出力パケットとして選択して出力する。

【0124】このような処理によって、各々のフローに対して、その要求するサービス品質(QoS)に応じ、 40

低ジッタ保証、低遅延保証を行なうとともに、フロー間で公平にかつ出力リンクを最大に有効利用してパケット転送を行なうことができる。

【0125】

【発明の効果】以上説明したように、本発明によれば、バックログキューにより使用される予約バンド幅の合計値の変化によって転送終了予定時刻が変動する各キューの先頭パケットに対して、転送終了予定時刻の最も早い先頭パケットを選択して転送する重み付け公平キューイング(WFQ)の処理において、バックログキューの予約バンド幅の合計値の変動に対する該転送終了予定時刻の相対的な順位変動の依存性の低い先頭パケット同士をクラスタ化し、過去に計算した各クラスタ内の出力優先順位をそのまま再利用して、各クラスタの第1順位パケットのみに対して転送終了予定時刻の計算を行ない、最優先パケットを選出することにより、重み付け公平キューイング(WFQ)処理における計算量を大幅に削減することができる。

【0126】また、転送終了予定時刻の計算において、時間軸が実時間軸であるため、複数のスケジューラで個々に選出された第1優先のパケットに対する選択処理が容易となり、複数のスケジューラによる選択処理を容易に連携することができる。

【0127】

【図面の簡単な説明】

【図1】本発明のパケット転送制御の原理説明図である。

【図2】転送終了予定時刻 $F_{i,1}$ と予約バンド幅使用率 γ との関係を表したグラフである。

【図3】本発明のパケット転送制御装置の構成を示す図である。

【図4】本発明における複数のスケジューラによるパケット転送制御の説明図である。

【図5】重み付け公平キューイング(WFQ)の原理説明図である。

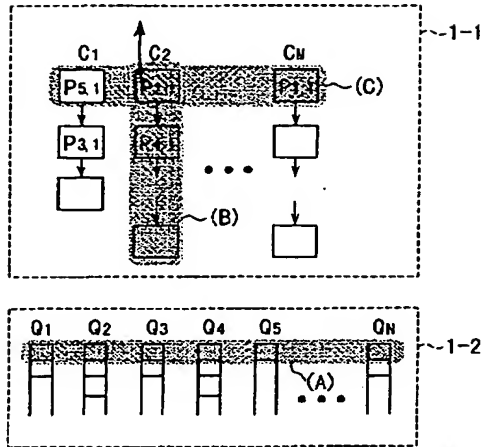
【符号の説明】

1-1 パケットをクラスタ化して転送制御するスケジューラ

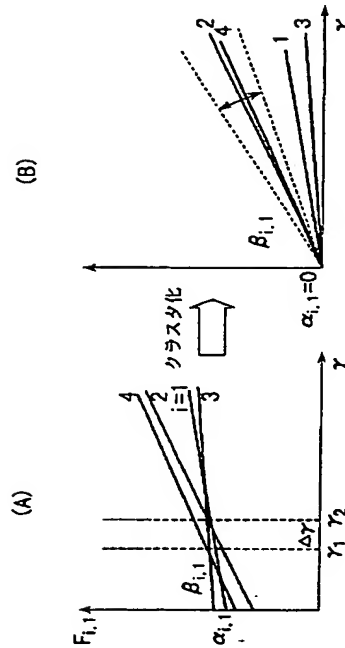
1-2 各フロー対応にパケットを格納するN個のキュー($Q_1 \sim Q_N$)から成るマルチキュー

【図1】

本発明のバケット転送制御の原理説明図

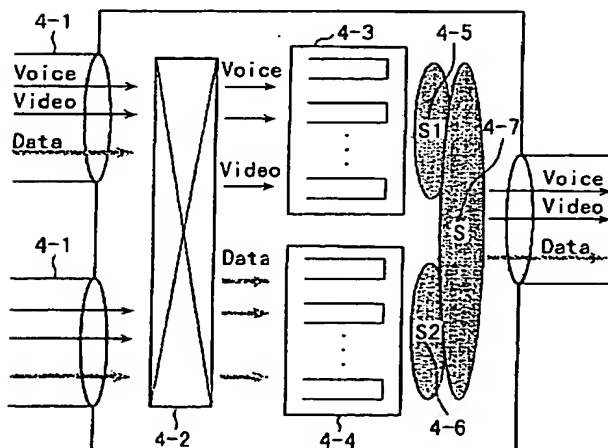


【図2】

転送終了予定時刻 $F_{i,1}$ と予約バンド幅使用率 γ との関係を表したグラフ

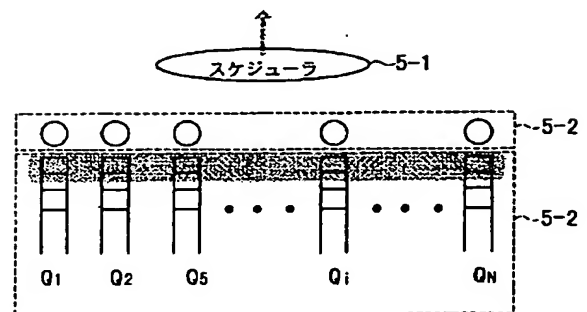
【図4】

本発明における複数のスケジューラによるバケット転送制御の説明図



【図5】

重み付け公平キューイング(WFQ)の原理説明図



【図3】

本発明のバケット転送制御装置の構成を示す図

